13th International Conference on Sustainable Energy technologies (SET2014)
25-28th August, 2014
Geneva
Paper ID: SET2014-E40137

Real-Time Optimization: Optimizing the Operation of Energy Systems in the Presence of Uncertainty and Disturbances.

Grégory François[1,2*] and Dominique Bonvin[1]

[1] Laboratoire d'Automatique, Ecole Polytechnique Fédérale de Lausanne, Switzerland
[2] University of Applied Sciences and Arts, Western Switzerland, Delémont, Switzerland

[*] *Corresponding email: gregory.francois@hes-so.ch*

ABSTRACT
In practice, the quest for the optimal operation of energy systems is complicated by the simultaneous presence of operating constraints, among which the need for producing the power required by the user, and of uncertainty. The latter concept incorporates the potential inaccuracies of the models at hand but also degradation effects or unexpected changes, such as, e.g. random load changes or variations of the availability of the energy source for renewable energy systems. Since these changes affect the optimal values of the operating conditions, online adaptation is required to ensure that the system is always operated optimally. This typically implies the online solving of an optimization problem. Unfortunately, the applicability and the performances of most model-based optimization methods rely on the quality of the available model of the system under investigation. On the other hand, Real-Time Optimization (RTO) methods use the available online measurements in the optimization framework and are, thus, capable of bringing the desired self-optimizing control reaction. In this article, we show the benefits of using several RTO methods (co-) developed by the authors to energy systems through the successful application of (i) "Real-Time Optimization via Modifier Adaptation" to an experimental Solid Oxide Fuel Cells (SOFC) stack, of (ii) the recently released "SCFO-solver[1]" to an industrial SOFC stack, and of (iii) Dynamic RTO to a simulated tethered kite for renewable power production. It is shown how such problems can be formulated and solved and significant improvements of the performances of the three aforementioned energy systems are illustrated.

KEYWORDS: *Real-Time Optimization; Energy Systems; Renewable Energy; Model Mismatch; Uncertainty.*

1 INTRODUCTION

In a world of increasing competition and emerging energy crisis, operating energy systems optimally is a must. This is particularly true for the most innovative devices or for renewable energy systems, since many of these devices need to prove that their higher investment costs can be compensated by higher operating performances. Standard approaches for determining the optimal operating conditions generally consist in the solving of a model-based optimization problem, for which a model of the system has to be available. Many schemes are available in the literature that can provide the optimal values of the degrees of freedom for given model under given experimental conditions. However, most of these optimization schemes suffer from the difficulty of obtaining reliable models of energy systems. These models consist in sets of algebraic and/or differential equations that depend on parameters, which are generally identified with a certain level of confidence. Meanwhile, it is not rare that models incorporate structural mismatch with the plant. One could argue that this is indeed always the case since, e.g., all real dynamical systems are of infinite order, while all dynamical models are of finite dimensions. Also real processes are prone to real-time disturbances, degradation, and many other unexpected (and thus un-modeled) events/phenomena that change the optimal values of the degrees of freedom with time. Thus, solving a model-based optimization problem is not sufficient for real-life applications and leads to sub-optimal performances, or worse, to potential violations of the operating constraints.

The fact that the optimal operating conditions of any real process can be drifting or changing abruptly calls thus for methods that are capable of rejecting the effect of all the aforementioned sources of uncertainty, while enforcing the satisfaction of operating constraints. This is exactly the scope of Real-Time Optimization (RTO) methods. RTO methods are popular in the chemical engineering field where they are performed at an intermediate frequency between the fast decision-making control layers (advanced and regulatory control) and the slower layers of planning and scheduling. Probably the most intuitive and popular family of techniques in the industry is the two-stage [1] (aka "two-step") approaches (RTO-TS). These methods use the difference between the output measurements and the corresponding model predictions to adapt the model parameters, with the updated model being used to repeat the optimization. Although appealing, it has been shown that these techniques are capable of rejecting the effect of parametric uncertainty but not of structural plant-model mismatch [2]. Recently, it has been proposed to update the model differently. Instead of adjusting iteratively the model parameters, measurement-based correction terms are iteratively added to the cost and constraint functions of the optimization problem. The technique, labeled RTO via modifier adaptation (RTO-MA), forces the modeled cost and constraints to match the plant values [2, 3]. The main advantage of RTO-MA compared to the RTO-TS lies in its ability to converge to the true plant optimum, even in the presence of structural plant-model mismatch [2]. Both families are referred to as explicit methods as they rely on the repeated solving of model-based optimization, but they differ on how measurements are used to reject the effect of uncertainty and model mismatch. In contrast, implicit methods, such as self-optimizing

[1] The acronym SCFO in the name of the solver (SCFO-solver) stands for "sufficient conditions for feasibility and optimality". Unfortunately this acronym is very similar to SOFC, which is used for "Solid Oxide Fuel Cells".

control [4] and NCO tracking [5], propose to adjust the degrees of freedom on-line in a control-inspired manner, by, e.g., forcing the necessary conditions of optimality of the real system to be satisfied.

So far, RTO methods have been mainly applied, sometimes routinely, to industrial chemical processes. This is motivated by historical reasons: until recently online optimization techniques were only applicable to slow processes (such as chemical, biotechnological or pharmaceutical processes), due to the heavy computational load. This is no longer the case and generally speaking, optimization methods are now considered as viable technologies [6, 7]. Indeed there are many reasons to use RTO methods for improving the performances of industrial energy systems. In [8], the authors justify the use of such methods for the optimization of heat and power systems with the following arguments: (i) changes in steam demand, price of electric power (potentially fast and frequent according to grid demands in deregulated markets), (ii) changes in boiler and turbine efficiencies, (iii) fast dynamics of most energy systems allow the use of steady state models, (iv) the resulting steady state optimization problems can be solved with a period of the order of minutes instead of the typical hours. Also, the authors claim that the online use of steady state RTO methods can prevent from using an advanced control layer with e.g. model-predictive control. From the authors' viewpoint these four arguments valid, but another added value of some of RTO methods, like the ones used in this study, lies in their capacity to reject the effect of parametric uncertainty (e.g. (ii)), process disturbances (e.g. (i), (iii)) but also of structural plant-model mismatch in a unified, theoretically sound, framework, provided the user can put its optimization problem in an appropriate form.

In this paper, we illustrate further the potential benefits of RTO methods for the optimization of energy systems under changing conditions, in the presence of structural plant model-mismatch, parametric uncertainty and disturbances. Since, practically speaking, solving an optimization problem for uncertain or badly modeled processes might seem to be obscure, we first show how these problems are formulated in the general setting and for application of the specific RTO approaches discussed hereafter. Then, we show how this can be done in the context of energy systems in application-oriented manner. Three methods and examples are considered: (i) RTO-MA and (ii) the recently released open-source "SCFO-solver" [8], and (iii) dynamic RTO. Successful applications are shown for a 6-cells experimental SOFC stack (i) and for a 2-cells industrial SOFC stack (ii). Also, successful application of dynamic RTO to a simulated tethered kite for power production, with ongoing experimental application. Thus, one of the messages of this contribution is thus to show that the gap between mathematically sound optimization methods and real-life applications to energy systems is not as big as one may think.

The paper is organized as follows. The formulation of RTO problems is presented in Section 2. Section 3 discusses the examples and the corresponding results, while Section 4 concludes the paper.

2 REAL-TIME OPTIMIZATION

2.1 Problem Formulation

The main problem at hand is to optimize the steady-state performances of a real energy system (hereafter referred to as "the plant"):

$$\begin{aligned} \min_{\mathbf{u}} \quad & \phi_{\mathrm{p}}\left(\mathbf{u}\right) \\ \text{s.t.} \quad & \mathbf{G}_{\mathrm{p}}\left(\mathbf{u}\right) \leq \mathbf{0} \end{aligned}, \tag{1}$$

where ϕ_{p} is the plant cost (i.e. the performance indicator), $\mathbf{G}_{\mathrm{p}}$ the n_g-dimensional vector of plant constraints and $\mathbf{u}$ the m-dimensional vector of constant inputs (the so-called "degrees of freedom"). The necessary conditions of optimality (NCO) for the plant read [9]:

$$\begin{aligned} \mathbf{G}_{\mathrm{p}}(\mathbf{u}^*) \leq \mathbf{0}, \quad \boldsymbol{\nu}_{\mathrm{p}}^* \geq \mathbf{0}, \quad \boldsymbol{\nu}_{\mathrm{p}}^{*T}\mathbf{G}_{\mathrm{p}}(\mathbf{u}^*) = 0 \\ \nabla\phi_{\mathrm{p}}(\mathbf{u}^*) + \boldsymbol{\nu}_{\mathrm{p}}^{*T}\nabla\mathbf{G}_{\mathrm{p}}(\mathbf{u}^*) = \mathbf{0} \end{aligned}, \tag{2}$$

where ν_{p} is a vector of Lagrange multipliers. The conditions (2) are the necessary conditions that $\mathbf{u}$ needs to satisfy to be a candidate solution to Problem (1). Indeed most numerical solvers rely on algorithms that directly or indirectly try to enforce the satisfaction of the conditions (2). Model-based optimization solvers can be used whenever a steady-state model of the plant is available. For the sake of simplicity, we will assume in this article that the cost function and the constraints are directly measured and modeled. This assumption is in no way restrictive and is just made to avoid the introduction of too many notations. Thus, the model-based optimization problem reads:

$$\begin{aligned} \min_{\mathbf{u}} \quad & \phi\left(\mathbf{u},\boldsymbol{\theta}\right) \\ \text{s.t.} \quad & \mathbf{G}\left(\mathbf{u},\boldsymbol{\theta}\right) \leq \mathbf{0} \end{aligned}, \tag{3}$$

where ϕ is the model cost, $\boldsymbol{\theta}$ the n_θ-dimensional vector of model parameters, and $\mathbf{G}$ the n_g-dimensional vector of model constraints. The model NCO read [9]:

$$\begin{aligned} \mathbf{G}\left(\mathbf{u}^*,\boldsymbol{\theta}\right) \le \mathbf{0}, \quad \boldsymbol{\nu}^* \ge \mathbf{0}, \quad \boldsymbol{\nu}^{*T}\mathbf{G}\left(\mathbf{u}^*,\boldsymbol{\theta}\right) = 0 \\ \nabla\phi\left(\mathbf{u}^*,\boldsymbol{\theta}\right) + \boldsymbol{\nu}^{*T}\nabla\mathbf{G}\left(\mathbf{u}^*,\boldsymbol{\theta}\right) = \mathbf{0} \end{aligned}, \tag{4}$$

where $\boldsymbol{\nu}$ is the vector of Lagrange multipliers for the model. This is where the discrepancies between the model and the plant become problematic. As mentioned above, to be optimal for the plant, the optimal inputs $\mathbf{u}^*$must satisfy the NCO (2). However, solving the model-based optimization problem (3), guarantees only that the conditions (4) are satisfied. Indeed, in the presence of plant-model mismatch and disturbances, the solution of Problems (1) and (3) will typically differ, with the solution of (4) not satisfying (2).

2.2 Philosophy of RTO methods

As said in the introduction, what makes RTO methods capable of rejecting the effect of uncertainty is the use of measurements and their appropriate use. If the cost function and the constraints of the plant are measured (which is typically the case) it is readily seen, whenever it occurs, that their values differ from their model predictions at any given $\mathbf{u}$. This difference between the predictions and the corresponding measurements is indeed the common driving force of most RTO methods.
The way RTO methods use measurement is depicted in Figure 1 and is put in perspective with the standard framework of process optimization.

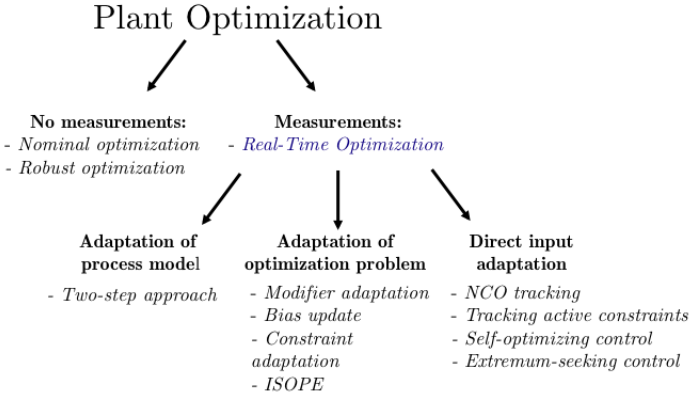

Fig 1. Classification of Several RTO methods and of their philosophies

As indicated by Fig.1 there are three main ways of using the available process measurements:

1. Measurements can be used to refine the model. By model refinement, it has to be understood that the mathematical expressions of the functions ϕ and $\mathbf{G}$ are updated in between consecutive optimization iterations. This adaptation is performed by means of the update of the model parameters $\boldsymbol{\theta}$. Two problems are iteratively solved: the identification problem and the optimization problem (3), for which ϕ and $\mathbf{G}$ are modified at each iteration.
2. Measurements can also be used to modify directly the optimization problem. With these methods, the driving force is used for the determination of corrective terms that are directly added to the expressions of ϕ and $\mathbf{G}$ in the formulation (3), which are kept unchanged. No parameter identification is performed, but a modified version of the optimization problem (3) is solved, *at each iteration*.
3. The third class of methods proposes to use the driving force to directly modify the inputs $\mathbf{u}$, in a control-inspired manner. With these methods, the problem (3) is only solved once, for the initialization of $\mathbf{u}$. The main difficulty is to construct a control problem, which has the desired self-optimizing properties, i.e. which guarantees that the resolution of the control problem implicitly enforces the resolution of Problem (1).

2.3 Real-Time Optimization via Modifier Adaptation (RTO-MA)

RTO-MA belongs to the second column of RTO methods depicted in Fig. 1. Measurements, or in better words, deviations between predicted and measured quantities are used to modify the model-based problem formulation by adding affine-in-input correction terms to ϕ and $\mathbf{G}$.

With RTO-MA, a modified version of Problem (3) is solved repeatedly at steady state, until convergence to the plant optimum is reached. For example, at the k^{th} iteration, the following problem is solved:

$$\begin{aligned} \mathbf{u}_{k+1}^* = \arg\min_{\mathbf{u}} \quad & \phi_m\left(\mathbf{u},\boldsymbol{\theta}\right) := \phi\left(\mathbf{u},\boldsymbol{\theta}\right) + \boldsymbol{\Lambda}_k^{\phi^T}\left(\mathbf{u}-\mathbf{u}_k^*\right) \\ \text{s.t.} \quad & \mathbf{G}_m\left(\mathbf{u},\boldsymbol{\theta}\right) := \mathbf{G}\left(\mathbf{u},\boldsymbol{\theta}\right) + \boldsymbol{\varepsilon}_k + \boldsymbol{\Lambda}_k^{G^T}\left(\mathbf{u}-\mathbf{u}_k^*\right) \le \mathbf{0} \end{aligned}, \tag{5}$$

where $\mathbf{u}_k^*$ denotes the solution at iteration k that will be used at the next iteration. Optimization problem (3) is modified by the addition of a linear correction term to the cost and an affine correction term to the constraints, with the modifiers defined as follows $\boldsymbol{\varepsilon}_k \coloneqq \mathbf{G}_p(\mathbf{u}_k^*) - \mathbf{G}(\mathbf{u}_k^*, \boldsymbol{\theta})$, $\boldsymbol{\Lambda}_k^{\boldsymbol{\phi}} \coloneqq \nabla\phi_p(\mathbf{u}_k^*) - \nabla\phi(\mathbf{u}_k^*, \boldsymbol{\theta})$ and $\boldsymbol{\Lambda}_k^G \coloneqq \nabla\mathbf{G}_p(\mathbf{u}_k^*) - \nabla\mathbf{G}(\mathbf{u}_k^*, \boldsymbol{\theta})$. The nicest feature of RTO-MA lies in that, upon convergence, (7) and (1) share the same NCO, that is, convergence to the plant optimum is possible despite the presence of uncertainty [2]. The corresponding converged modifiers are $\boldsymbol{\varepsilon}_\infty$, $\boldsymbol{\Lambda}_\infty^{\boldsymbol{\phi}}$ and $\boldsymbol{\Lambda}_\infty^G$. What it makes become handy lies in the fact that, with the so-called 1st-order modifications, the gradients of the modified model cost and constraints are corrected $\boldsymbol{\phi}_m$ and $\mathbf{G}_m$, although the functions $\boldsymbol{\phi}$ and $\mathbf{G}$ are not modified directly, contrary to what happens with the two-step approaches.

As already mentioned, with the two-step approaches, the mathematical expression of the model is modified, by means of the update of the model parameters $\boldsymbol{\theta}$. This identification is performed with the objective of minimizing the distance between the predicted and the measured values of $\boldsymbol{\phi}$ and $\mathbf{G}$, which ultimately leads to perfect fitting of the data collected. But nothing guarantees that the gradients of the modified functions $\boldsymbol{\phi}$ and $\mathbf{G}$ will match the gradients of $\boldsymbol{\phi}_p$ and $\mathbf{G}_p$ as fitting the data (the measured outputs) does not necessarily mean matching the necessary conditions of optimality.

One sees clearly from (2) and (4) that the conditions of optimality are of 1st order. Hence, although it is needed that the values of the predicted constraints match the values of the plant constraints to ensure feasibility, the matching of the gradients is required for enforcing the satisfaction of the sensitivity conditions. While the two-step approaches cannot guarantee this latter point, RTO-MA is by construction capable of doing so. Of course, there are conditions for RTO-MA to be capable of converging to the plant optimal operating conditions. This is a property that the original model $\boldsymbol{\phi}$ and $\mathbf{G}$ needs to have, which is referred to as in the literature as model adequacy [2]. This is obvious since without such conditions, this would mean that any process could be optimized with any model. These conditions also exist for the two-stage approaches [10] and it hopefully turns out that the adequacy conditions for RTO-MA are much less restrictive that the ones for the two-step approaches. It has also been shown recently that some of these conditions could always been matched when using convex approximations of the modeled functions $\boldsymbol{\phi}$ and $\mathbf{G}$ [11]. In addition recent work on gradient estimation techniques [12, 13, 14] has compared or proposed methods that can be used for estimating the 1st-order modifier terms, denoted in this article by the Greek letter $\boldsymbol{\Lambda}$.

2.4 The SCFO-solver

The SCFO-solver is a recently-released open-source solver [17] that implements the so-called "sufficient conditions for feasibility and optimality" (SCFOs). These conditions, which are given below, are sufficient conditions, which when satisfied, ensure that a sequence of experiments converge to a local optimal solution of Problem (1), with the additional property that each iterate is feasible.

$$
\begin{aligned}
& g_{p,j}(\mathbf{u}_k) + \sum_{i=1}^{n_u} \kappa_{p,ji} \left| u_{k+1,i} - u_{k,i} \right| \le 0, \quad \forall j = 1,\ldots,n_{g_p} \\
& g_j(\mathbf{u}_{k+1}) \le 0, \qquad\qquad\qquad\qquad \forall j = 1,\ldots,n_g \\
& \mathbf{u}^L \le \mathbf{u}_{k+1} \le \mathbf{u}^U \\
& \nabla g_{p,j}(\mathbf{u}_k)^T (\mathbf{u}_{k+1} - \mathbf{u}_k) < 0, \quad \forall j : g_{p,j}(\mathbf{u}_k) \approx 0 \\
& \nabla g_j(\mathbf{u}_k)^T (\mathbf{u}_{k+1} - \mathbf{u}_k) < 0, \quad \forall j : g_j(\mathbf{u}_k) \approx 0 \\
& \nabla \phi_p(\mathbf{u}_k)^T (\mathbf{u}_{k+1} - \mathbf{u}_k) < 0 \\
& \nabla \phi_p(\mathbf{u}_k)^T (\mathbf{u}_{k+1} - \mathbf{u}_k) + \frac{1}{2} \sum_{i_1=1}^{n_u} \sum_{i_2=1}^{n_u} M_{\phi,i_1 i_2} \left| (u_{k+1,i_1} - u_{k,i_1})(u_{k+1,i_2} - u_{k,i_2}) \right| \le 0
\end{aligned}
\quad , \tag{6}
$$

In (6) the SCFOs for the real process are summarized. Two kinds of constraints are considered (with or without the subscript "p"). to account for the situations where there are simultaneously accurately and inaccurately (with subscript "p") modeled constraints. Note that these conditions are not implementable per se, as they are mainly of mathematical nature. They rely on the assumption that all functions are C^2 and Lipschitz-continuous [15], with estimates of the Lipschitz constants ($\boldsymbol{\kappa}$) and of the quadratic upper bound (on the cost, here the M_ϕ), available. The proof that these conditions are indeed sufficient is very much involving and requires several intermediate lemmas and theorem. The presentation or the discussion of these results is beyond the scope of this article and the interested reader is kindly invited to read the following reference [16].

Although very technical, these conditions can still be interpreted. The three first equations enforce feasibility at the following iterate: since the global maximal amplitude of the variations of all the process constraints is supposed to be known (through Lipschitz constants), a simple reverse engineering analysis allows to bound the maximum distance between the current and the

future set of operating conditions such that no constraint will be violated. For example, if one of the constraints is on temperature (say limited by 100 degrees), if the current temperature is at 80 degrees, and if the Lipschitz constant of this constraint is equal to 4, this means that the subsequent iterate must lie in a ball of radius 5 around the current iterate to guarantee that no constraint violation will occur. On the other hand, the two last equations ensure that the plant cost will be better at the subsequent iterate (see [16] for all mathematical details). Unfortunately going in the good direction with feasible iterates is not sufficient as we need also to avoid infinitely small step sizes. This is enforced by the 4th and the 5th equations of (6), the so-called "projection conditions" where it is guaranteed that the next iterate lies strictly in the set of local descent directions of all constraints.

As such these conditions are not implementable, mainly because of the presence of the strict inequalities. The aim of the solver is thus to formulate these conditions in an implementable way and, in turn, to allow their stand-alone implementation or in connection to the chosen RTO method. Without entering the details of [16], the solver [17] becomes implementable when the strict inequalities are relaxed by means of individual back-offs, backed-off broad inequalities replacing thus the original strict inequalities. These slack variables are indeed the tuning parameters of the solver, and it is shown in [16] that by asymptotically reducing the values of these variables, the user will converge in a finite number of iterates arbitrarily close to a local optimum of (1). Indeed, this solver can thus be seen itself as a stand-alone RTO method or, alternatively, as a framework for RTO methods. With the latter, the set of operating conditions obtained with a given RTO solver will follow the modifications induced by the implementable version of (6), guaranteeing thus feasibility and cost-improving iterates. The user can thus benefit from the power of the most recent model-based RTO algorithms (speed of convergence, low failure rates) while reducing the risks of constraint violations or of sub-optimality induced by the use of potentially inaccurate models.

3 RESULTS AND DISCUSSION

3.1 RTO-MA of an experimental SOFC stack

The first case study considers the optimization of the electrical efficiency of a Solid Oxide Fuel short-stack developed at EPFL for HTceramix – SOFCpower (http://www.sofcpower.com/48/company-profile.html) [18, 19].

This 6-cells stack consists of planar anode-supported cells with an active area of 50 cm^2, pressed between gas-diffusion layers (SOFConnex^{TM}) and metallic interconnector plates. The anodes are made of nickel/yttrium stabilized-zirconia (Ni-YSZ) cermet, while the electrolyte consists of dense YSZ. The cathodes are made of screen-printed (La, Sr)(Co, Fe)O_3, allowing standard operation temperatures between 650◦C and 850◦C. A photo of a typical assembly is given on the left plot Fig. 2, together with a scheme of the device on the right hand side.

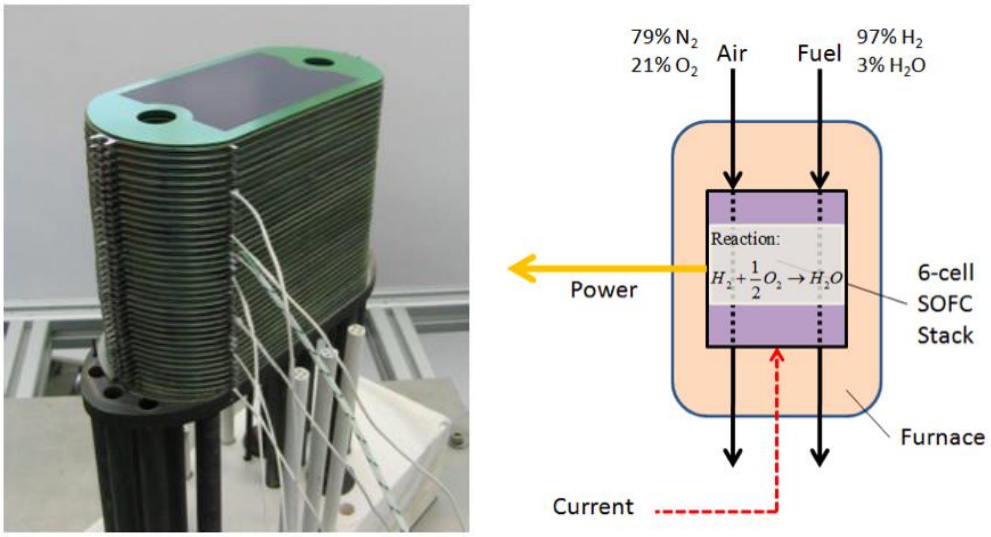

Fig 2. The experimental SOFC stack

A steady-state simplified model of this device is available [20]. It is known that this model, although it is a reasonable tendency model of the real stack, is not very accurate in terms of predictions of the polarization curve, the cell potential and the power produced. This is why a standard model-based optimization approach has not been considered to be sufficient [21].

The experimental setup has 3 degrees of freedom (manipulated inputs): the fluxes of Hydrogen and Oxygen and the current intensity, and three outputs of interest: the power density, the cell potential and the electrical efficiency. The aim is to maximize the electrical efficiency while: (i) producing the (unknown and varying) power density asked by the user, (ii) maintaining the cell potential above a lower bound of 0.75V, (iii) maintaining the fuel utilization below 0.75, and (iv) maintaining the air ratio between 4 and 7. This problem can formulated as a plant optimization problem:

$$\begin{aligned}
\max_{\mathbf{u}} \eta_p(\mathbf{u}) &:= \frac{U_{cell,p} N_{cells} I p_{el,p}}{\dot{n}_{H_2} Q_L} \\
\text{s.t.} \quad & p_{el,p}(\mathbf{u}) = p_{el}^S \\
& U_{cell,p}(\mathbf{u}) \geq 0{,}75\text{V} \\
& \nu(\mathbf{u}) := \frac{N_{cells} I}{2\dot{n}_{H_2} F} \leq 0{,}75 \\
& 4 \leq \lambda_{air}(\mathbf{u}) := 2\frac{\dot{n}_{O_2}}{\dot{n}_{H_2}} \leq 7 \\
& u_1 \geq 3.14\text{ml/(min.cm}^2\text{)} \\
& u_3 \leq 30\text{A}
\end{aligned} \tag{7}$$

Where $\mathbf{u} = \left[\dot{n}_{O_2},\ \dot{n}_{H_2},\ I\right]^T$ is the vector of controlled inputs. Without input errors, these degrees of freedom are the same for the model.
The vector of inputs is thus composed of the fluxes of oxygen $\dot{n}_{O_2}$ and hydrogen $\dot{n}_{H_2}$, and of the current I. $U_{cell,p}$, N_{cells}, $p_{el,p}$, and Q_L denote the cell potential, the number of cells, the power density delivered by the stack and the lower heating value of the fuel. N_{cells} and Q_L are known with certainty, but this is not the case of the two others, justifying thus the subscript "p" for the electrical efficiency. The first equality constraint $p_{el,p} = p_{el}^S$, means that the power produced should match in real-time the required value. This corresponds to a major challenge in that p_{el}^S is never known in advance. The three next inequality constraints mean that the cell potential $U_{cell,p}$ is lower bounded, the fuel utilization $\nu(\mathbf{u})$ (where F denotes the Faraday constant) is upper bounded, while the air ratio should remain between two bounds. Note here that the cell potential is not known with certainty. It is indeed modeled on the basis an equivalent-circuit approach [22], but it is widely admitted that, similarly to the power density produced by the real stack, model predictions are not, by far, perfect. On the other hand, the fuel utilization and the air ratio depend only on the value of the inputs, which are known with certainty. This justifies the use of the subscript "p" for the cell potential and its omission for the two other quantities. Finally, the two last inequalities are just classical main bounds on the input variables (e.g. maximum flowrates).

Thus, the optimization problem of Eqn. (7) suffers from a certain level of model-mismatch, mainly because the cell potential and the power produced are not well predicted by the model at hand. Therefore, using a standard model-based optimization method is very unlikely to provide the user with the plant optimal inputs. Anyway, the fact that the load p_{el}^S changes with time, calls for an update of the inputs applied to the real stack for maintaining maximal electrical efficiency, regardless these inputs have been ideally initialized or not. This justifies the use of either RTO-MA or of the SCFO solver, since these two approaches/tools are capable of rejecting the effect of uncertainty and thus, tracking the changes in the set of optimal operating conditions.

Without entering the details of [21], RTO-MA was first applied in its simplest form, i.e. only the 0^{th} –order modifier terms (the epsilons in (5)) were used. This is justified since it has been shown experimentally that the solution is almost always determined by the constraints. At each re-optimization stage k, the following optimization problem was solved:

$$\begin{aligned}
\mathbf{u}_{k+1}^* &:= \arg\max\left(\eta_m := \eta(\mathbf{u}) + \varepsilon_k^{\eta}\right) \\
\text{s.t.} \quad & p_{el,m}(\mathbf{u}) := p_{el}(\mathbf{u}) + \varepsilon_k^{p_{el}} = p_{el}^S \\
& U_{cell,m}(\mathbf{u}) := U_{cell}(\mathbf{u}) + \varepsilon_k^{U_{cell}} \geq 0{,}75\text{V} \\
& \nu(\mathbf{u}) \leq 0{,}75 \\
& 4 \leq \lambda_{air}(\mathbf{u}) \leq 7 \\
& u_1 \geq 3.14\text{ml/(min.cm}^2\text{)} \\
& u_3 \leq 30\text{A}
\end{aligned} \quad , \tag{8}$$

with $\eta(\mathbf{u})$, $p_{el}(\mathbf{u})$ and $U_{cell}(\mathbf{u})$ being the values of the electrical efficiency, of the power density and of the cell potential predicted by the model of [20]. The modifier terms correspond to the difference between predicted and measured quantities at the previous iteration. Also, a unit gain exponential filter is applied to avoid too abrupt modifications of the optimization problem. This filter introduces thus, through its gains K, parameters that can be tuned to facilitate convergence to the plant optimum [20].
The modifier computation reads:

$$\begin{aligned}
\varepsilon_k^{\eta} &= \left(1 - K_{\eta}\right)\varepsilon_{k-1}^{\eta} + K_{\eta}\left(\eta_{\mathbf{p}}\left(\mathbf{u}_{k-1}^*\right) - \eta\left(\mathbf{u}_{k-1}^*\right)\right) \\
\varepsilon_k^{p_{el}} &= \left(1 - K_{p_{el}}\right)\varepsilon_{k-1}^{p_{el}} + K_{p_{el}}\left(p_{el,p}\left(\mathbf{u}_{k-1}^*\right) - p_{el}\left(\mathbf{u}_{k-1}^*\right)\right) \\
\varepsilon_k^{U_{cell}} &= \left(1 - K_{U_{cell}}\right)\varepsilon_{k-1}^{U_{cell}} + K_{U_{cell}}\left(U_{cell,p}\left(\mathbf{u}_{k-1}^*\right) - U_{cell}\left(\mathbf{u}_{k-1}^*\right)\right)
\end{aligned} \quad , \tag{9}$$

Eqn (9) indicates that the 0[th]-order modifiers are indeed the filtered differences between the three measured values of the efficiency, of the power density and of the cell potential and the three corresponding predicted values, at the previous set of optimal operating conditions. Upon convergence, whenk → ∞, it is straightforward to show that $\eta(\mathbf{u}_\infty^*) = \eta_m(\mathbf{u}_\infty^*)$, $p_{el,p}(\mathbf{u}_\infty^*) = p_{el,m}(\mathbf{u}_\infty^*)$, and $U_{cell,p}(\mathbf{u}_\infty^*) = U_{cell,m}(\mathbf{u}_\infty^*)$, i.e. that the modified modeled constraints match the real constraints at the converged inputs. In other words, with the modifiers (in theory with both 0[th] and 1[st]-order modifier terms), the necessary conditions of optimality of the modified model-based problem match those of the real problem, and so do the optimal solutions.

The following figure depicts the implementation of the algorithm to the experimental stack.

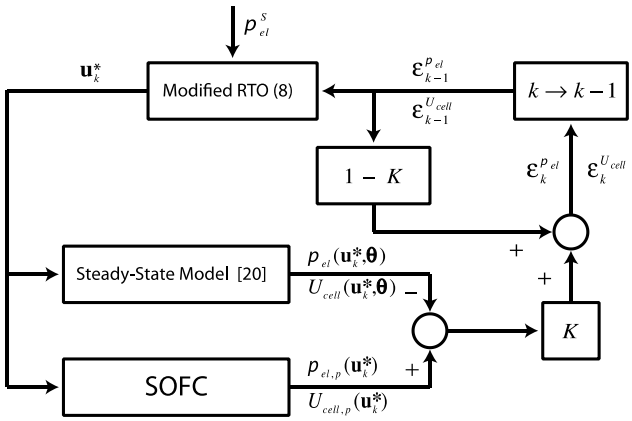

Fig 3. RTO-MA applied to the experimental SOFC stack. The modified problem of (8) is solved for the current value of p_{el}^S and the obtained optimal inputs are applied to the steady-state model and to the stack in parallel. Due to model-mismatch and parameteric uncertainty (inaccurate values of the model parameters $\boldsymbol{\theta}$), the measured values differ from the model predictions. These differences are filetered (top-right blocks) to compute the new value of the modifiers to implement in the modified problem of Equation (8)

The implementation of the scheme of Fig. 3 to the 6-cells stack of Fig. 2 was performed under the following scenarios [21]:

1. p_{el}^S varies from 0.30 W/cm^2 to 0.38 W/cm^2 then back to 0.30 W/cm^2 every 90 minutes
2. p_{el}^S varies randomly every 5 minutes between 0.30 and 0.38 W/cm^2

For scenario 1, RTO-MA of (8) was solved every 30 minutes, a period that is sufficiently long to guarantee that the real stack has reached steady state, and thus that steady-state measurements are compared to predictions at steady state. For scenario 2, RTO-MA of (8) was solved every 10 seconds. Here, steady state predictions are compared to transient measurements, since the dominant constant time of the temperature dynamics is about 30 minutes. This introduces additional model-mismatch in that the steady-state model is compared to transient quantities. For both scenarios, the filter gains were also varied, and the information regarding the load changes was unknown at the implementation level – it was just used for simulating the behavior of the user [21]. Fig. 4 depicts the results obtained for scenario 2.

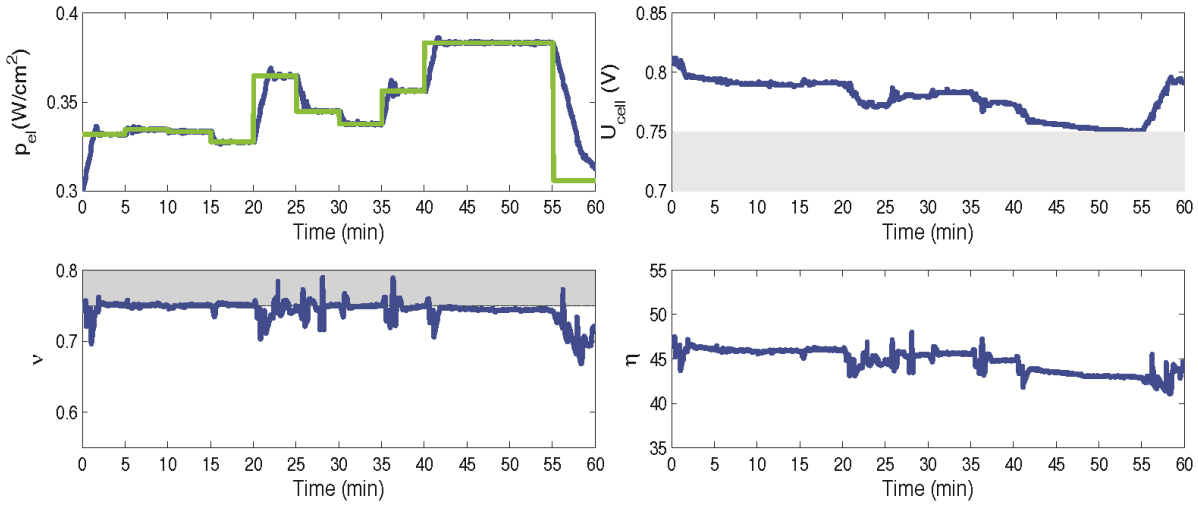

Fig 4. Experimental results of the application of RTO-MA to the experimental SOFC stack. Grey areas depict forbidden zones for either the cell potential (top-right) or for fuel utilization (bottom left). Solid blue lines depict measurements on the real stack, while the green curve is the unknown, randomly changing, load.

It is readily seen from Fig. 4 that the application of RTO-MA to the experimental SOFC stack is capable of (i) ensuring a fast tracking of the random (unknown) load changes (see the top left plot, where the blue curve follows the green one), (ii) maintaining the cell potential and the fuel utilization in the correct side of the constraints (top right, and bottom left, respectively), although marginal constraint violations are observed for fuel utilization that are due to the poor tuning of low level controllers. Finally (iii), the bottom right depicts the electrical efficiency. It is hard to quantify how optimal the performances are, since the optimal solution is not known *a priori*, but it is clear that the electrical efficiency is maintained at high values despite to trend of the random profile to increase the load. Our industrial partner validated these efficiencies to be close to the expected maximal performances, for the given loads. This is indeed a very good and promising result [21], which allowed us to go one step further and apply the SCFO-solver to, this time, an industrial stack.

3.2 Application of the SFCO-solver to an industrial SOFC stack

The second case study discussed hereafter considers the optimization of the electrical efficiency of a two-cells SOFC stack for HTceramix - SOFCpower. The presentation here will be very short, as the results presented in this section result from the direct application of the SCFO-solver to the problem (7). Note that the method described in [12] has been used for estimating the gradients on the real stack, while the solver has been coupled to a rudimentary gradient-descent with variable step size RTO algorithm. To simplify the analysis, the case where the solver is used for optimizing the electrical efficiency at constant load is presented. These are preliminary, unpublished results, but they were generated over a very short-time frame, which is a good indication of the applicability of our solver [17] to industrial facilities.

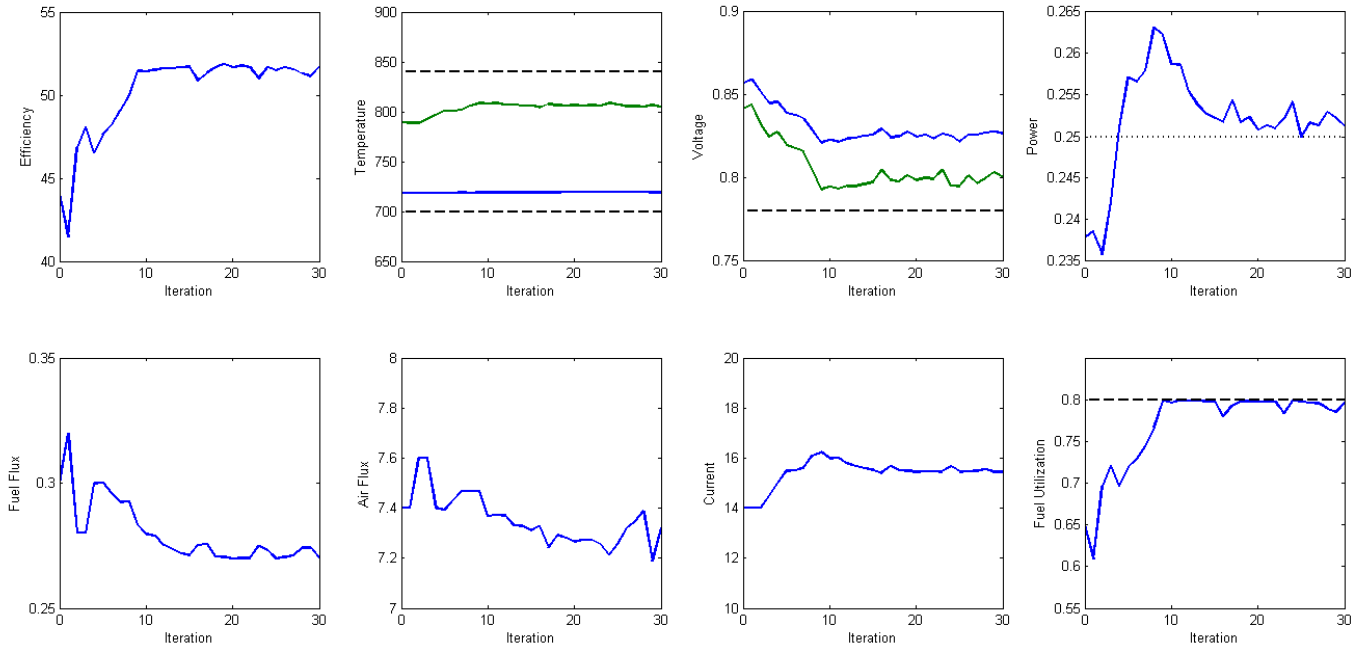

Fig 5. Industrial application of the SCFO-solver to a 2-cells SOFC stack.

Inspecting Fig. 5, where the results of the application of the solver to the industrial stack are depicetd, it is clear that the solver achieves all its goals. Starting from an infeasible point, in that the power load of 0.25 is not delivered (1st row, 4th plot), the power produced is first maximized, until the power demand is reached. Meanwhile, one sees that the electrical efficiency jumps from 42% at an initial infeasible point to nearly 53% at a feasible set of operating conditions (1st row, 1st plot). Also, the constraints on inlet and outlet temperatures (1st row, 2nd plot), on the two cell potentials (1st row, 3rd plot) and on fuel utilization (2nd row, 4th plot) are strictly satisfied at any iteration. At the time when we write this contribution, the solver has been (freely) downloaded nearly 10'000 times. We hope to receive feedback from the users, but, nevertheless, these preliminary results are very encouraging, and show a significant improvement in performances for an industrial (pilot-plant scale) energy system.

3.3 Dynamic RTO of a simulated Kite System

The idea of using kites as a renewable energy sources has received growing attention in recent years, both from industry and academia. Kites can fly at speeds many times that of the wind. At such high speeds a large aerodynamic force acts on the kite, and the force is transmitted to the ground via the cable. A number of different ways of using kites as an energy source are being investigated by 'kite-power' start-ups (Skysails GmBH, http://www.skysails.info; Makani Power, http://www.makanipower.com; Ampyx Power, http://www.ampyxpower.com). Two main reasons explain this recent increase in popularity: (i) the cable can transmit the power directly to the ground (without a tower, as with wind turbines) and (ii) it is much cheaper to use long cables and the access to the stable and strong high-altitude winds with a tethered kite than with a wind turbine.

For example, if the cable is connected to a generator, two phases can be envisaged:

1. Reel-Out: the tension in the cable forcibly unwinds the reel, the *generator produces electricity*.
2. Reel-In: *the generator works as a motor* to reel the cable back in.

The use of tethered kites for producing power is however still a research field and, for industrial applications, a niche market. Optimizing the operation of the kites is thus highly desirable, as these devices must still prove their efficiency to convince a broader audience of their potential.

For the example before, this means designing a trajectory that maximizes the cable tension during phase 1, while minimizing the tension during phase 2. Like in the previous sections, the standard approach would be to formulate a model-based problem, while it is widely admitted that constructing an accurate dynamic model of a kite is still an open problem. Several models have been proposed [23, 24, 25], with none being fully capable of accurately predicting a kite's behavior. Also, the kite's behavior is affected by time-varying disturbances, mainly the wind's speed and direction. These disturbances cannot be directly measured, as, due to wind shear, the wind at the kite's altitude is not the same as the wind at ground level. Hence, the need for a RTO method, but with the additional difficulty that a kite is fast, dynamic system, i.e. a dynamic RTO approach. Another limitation is that solving a dynamic optimization problem can take a significant amount of time, which can be incompatible with the real-time operation of such a fast system. Yet what makes it handy stems from the fact the optimization problem is indeed a periodic one. The kite should follow a trajectory in the air, which brings it back to its initial position. The aim of RTO-MA will thus to alter this trajectory in such a way that, from cycle to cycle, or from a series a tens of loops to another series of tens of loops, it rejects the effect of the disturbances to the performances of the kite.

The interested reader is invited to read [26] for the details regarding the formulation of the dynamic optimization problem. The goal is to maximize the average thrust, over one period (here corresponding to a series of loops), subject to a certain number of constraints (mainly here on the minimal altitude and on the fact that, at the end of the period, the position of the kite in space has to be equal to its initial position). For the application considered in this study (a kite used for pulling a boat, similarly to the products of Skysails GmBH, http://www.skysails.info), simulations with the simplified model of [26] showed that the solution of this optimization problem is composed of two arcs. The first arc is unconstrained while the second arc is one the height constraint, as illustrated by Fig. 6.

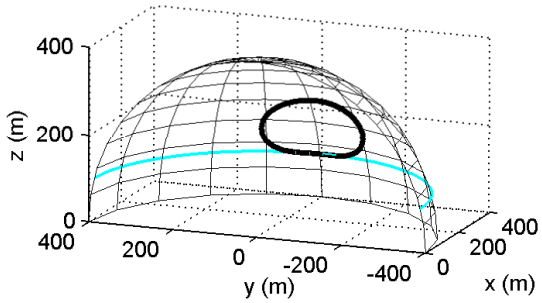

Fig. 6. Nominal optimal path on the sphere the kite is constrained to (the direction of flight is clockwise). The blue line is the height constraint.

Because of uncertainty, this trajectory can vary significantly, as illustrated by Fig. 7, where the evolution of the nominal path with regard to changes in the angle between the boat's direction and the wind relative to the boat is depicted.

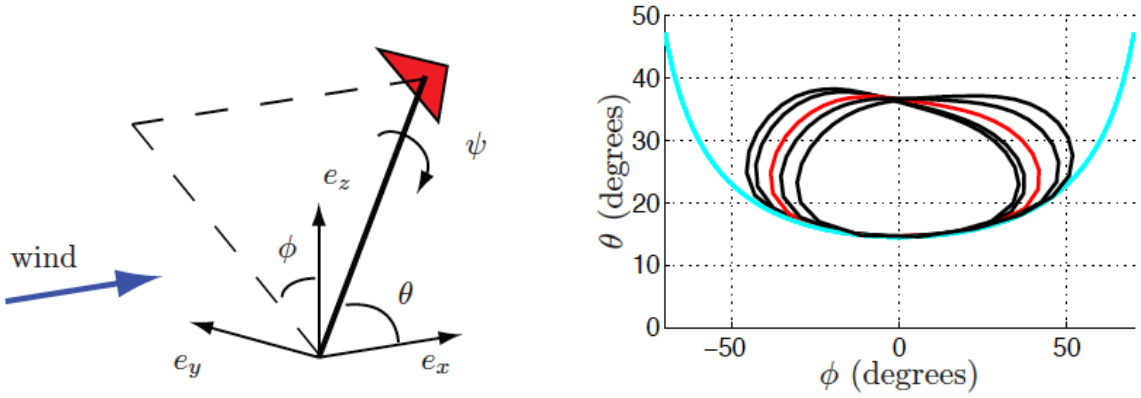

Fig. 7. System of spherical co-ordinates (left hand-side), with red triangle depicting the wing, and variation of the nominal path when the angle between the wind relative to the boat and the boat's velocity, varies from -15◦ to +15◦ around its modeled value

Fig. 8 depicts the scheme proposed in [26] to iteratively bring the required correction to the path that should follow the kite, in order to maximize the thrust.

This scheme is constructed based on the following mathematical reasoning:

- The solution to the model-based dynamic optimization problem is the optimal path $\mathbf{z}^*(\mathrm{l},\ \pi) \coloneqq [\theta^*(\mathrm{l},\ \pi),\ \phi^*(\mathrm{l},\ \pi)]^T$, where $\mathrm{l} \in [0,\ \mathrm{l_f}]$ is the length of the path and π the uncertain parameters in the model (denoted here with the Greek letter π because $\theta$is already used for an angle.

- This path is composed of two arcs : $\mathbf{z}^*(\mathrm{l},\ \pi) = \begin{cases} \mathbf{z}^*_{\mathrm{UC}}(\mathrm{l},\ \pi) \text{ for } \mathrm{l} \in [0, \mathrm{l_S}] \\ \mathrm{h} = \mathrm{h_{min}} \text{ for } \mathrm{l} \in [\mathrm{l_S}, \mathrm{l_f}] \end{cases}$, with $\mathbf{z}^*_{\mathrm{UC}}$ denoting the unconstrained arc discussed above, h and $\mathrm{h_{min}}$the altitude of the wing and its lower bound, and $\mathrm{l_S}$the point at which the constraint on altitude becomes active.

- In general, for parametric optimal control problems, small variations of the uncertain parameters π do not alter the sequence and type of intervals in the solution [27].

- Our simulation studies show that this is indeed true even for larger variations of π (as shown in Fig. 7).

- A Taylor series expansion of $\mathbf{z}^*$can be performed in the vicinity of π_0, the nominal value of the model parameters, which reads : $\mathbf{z}^*(\mathrm{l},\ \pi_0) = \begin{cases} \mathbf{z}^*_{\mathrm{UC}}(\mathrm{l},\ \pi_0) + \frac{\partial \mathbf{z}^*_{\mathrm{UC}}(\mathrm{l},\pi)}{\partial \pi}\Big]_{\pi_0} \Delta\pi \\ \mathrm{h} = \mathrm{h_{min}} \end{cases}$,
- this expansion yields the direction in which the path should be adapted to reject the effect of uncertainty on kite's performances, but does not provide with the amplitude of the adaptation.

We propose to determine this amplitude and in turn, adapt the path simultaneously, with the following RTO scheme [26].

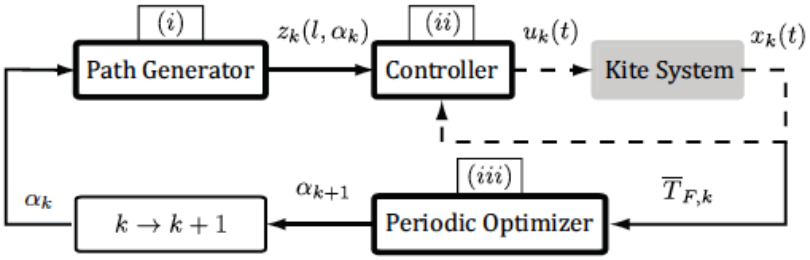

Fig. 8. The RTO scheme for the kite system

This scheme is composed of three blocks: (i) the path generator, which maps the amplitude of the adaptation to the change in the path to follow, (ii) the controller, which implements the tracking of the path (typically by means of a Model-Predictive Controller) and (iii) the periodic optimizer that adjusts the amplitude of the adaptation once per period, on the basis of the measurement of the average thrust over the k^{th} period (denoted $\overline{\mathrm{T}}_{\mathrm{F,k}}$).

In other words, the RTO algorithm (here a simple gradient descent method) uses $\overline{\mathrm{T}}_{\mathrm{F,k}}$ to compute the amplitude of modification α_{k} that is used to alter the path to follow, along the directions $\frac{\partial \mathbf{z}^*_{\mathrm{UC}}(\mathrm{l},\pi)}{\partial \pi}\Big]_{\pi_0}$, with the following update law:

$$\mathbf{z}_{\mathrm{k}}(\mathrm{l},\quad \alpha_{\mathrm{k}}) = \begin{cases} \mathbf{z}^*_{\mathrm{UC}}(\mathrm{l},\quad \pi_0) + \dfrac{\partial \mathbf{z}^*_{\mathrm{UC}}(\mathrm{l},\ \pi)}{\partial \pi}\Big]_{\pi_0} \alpha_{\mathrm{k}} \text{ for } 0 \le \mathrm{l} \le \mathrm{l_S} \\ \mathrm{h} = \mathrm{h_{min}} \text{ for } \mathrm{l_S} \le \mathrm{l} \le \mathrm{l_f} \end{cases}, \tag{10}$$

The application of (9) to the simulated kite system leads to the improvement depicted in Fig. 9.

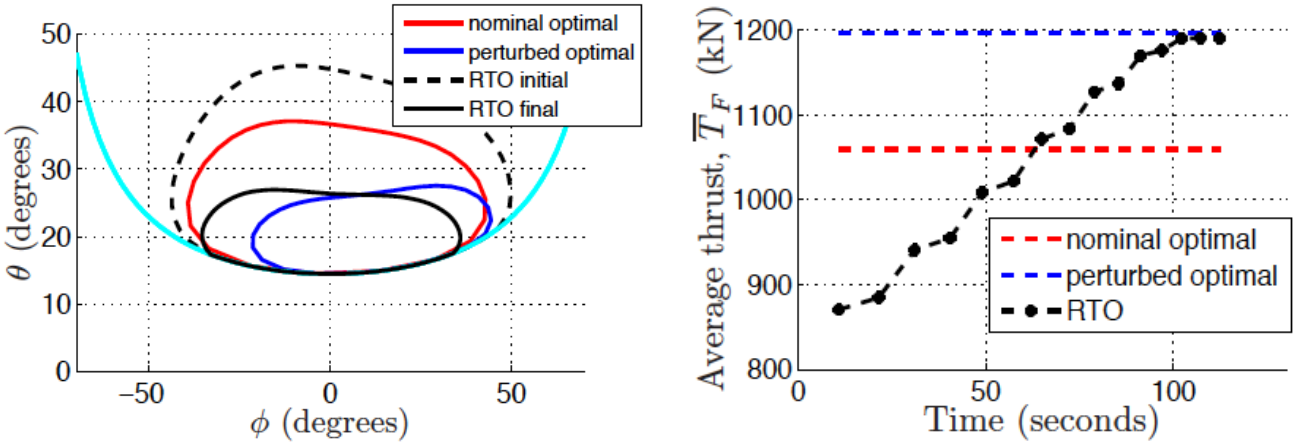

Fig. 9. Application of the RTO scheme to the simulated kite system.

On the left-hand side of Fig. 9, the evolution of the path from the initial trajectory (black dotted) to the converged trajectory (solid black) is plotted. Note that, the converged trajectory is almost the symmetric, but not the same than the true optimal one (solid blue). This is due to the fact that, in the vicinity of the optimal trajectory, a lot of trajectories yield to nearly the same performances. The right-hand side plot confirms this fact, as it is readily seen that the converged average thrust (black dotted) tends to the true solution (blue dotted). Note however, that the converged trajectory follows the "bean-shape" of the optimal solution. Overall, for this uncertainty scenario, the increase in performances reaches 25% over two minutes.

To conclude this section, we need to mention that we are currently developing an experimental device, a picture of which is given in Figure 10. The methodology will be applied to this experimental device in the weeks/months to come.

Fig. 10. Two pictures of the experimental kite system under development.

4 CONCLUSIONS

In this article, the potential benefits of using a selection of RTO methods to energy systems has been illustrated. It has been shown how these methods exploit the available measurements to compensate from model inaccuracies and to reject the effect of uncertainty and disturbances to optimize the performances of the real system at hand. In the field of energy, this is a highly desired property since, as it has been illustrated in this article, it allows to maintain optimality and constraint satisfaction despite changes of the load or of the source availability. The focus has been on the methods developed by the authors and especially on RTO-MA and on the SCFO-solver, with simulated, experimental and industrial examples.

The application of dynamic RTO and RTO-MA to a simulated kite system and to an experimental SOFC stack, respectively, has shown the potential of this technique to handle simultaneously parametric uncertainty, structural model-mismatch and disturbances in the context of variations of the availability of the power source or of the load, while improving the performances of the corresponding energy systems. Also, unpublished results concerning the succesful application of the SCFO-solver in an industrial context were presented, which show that these techniques can jump over the gap that still remains between the academic and the industrial worlds. But it is the opinion of the authors that the message of this paper is indeed broader than the sole focus on the apllication of well-chosen techniques or to well chosen systems.

The key message is that RTO methods and algorithms exist, which can be tailored to one's specific energy system. Some of these techniques are indeed (almost) ready to be used by practitioners for the online, in-situ improvement the operation of energy systems, even when the system at hand is hard to model. Meanwhile, the strength of the most recent methods is that they remain relatively easy to apply despite their strong theoretical fundations. As illustrated in this article, significant improvements can be obtained in terms of electrical efficiency of SOFC stacks, where it is made possible to track load changes at maximum electrical efficiencies. Meanwhile, on-going work on kites has shown that such methods can reject the effect of the changes in the availability of the source (here the direction and strength of the wind) for the most innovative renewable energy systems.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Gene Bunin and Mr. Sean Costello who contributed this work all the Master Students involved in the development of the Kite System, the SwissKitePower consortium and the HTCeramix-SOFC Power company.

REFERENCES

[1] T. Marlin and A. Hrymak. Real-Time Operations Optimization of Continuous Processes. In AIChE Symposium Series – CPC-V, 93 (1997), pp. 156-164..
[2] A. Marchetti, B. Chachuat and D. Bonvin. Modifier-Adaptation Methodology for Real-Time Optimization, Ind. Eng. Chem. Res., 48 (2009), pp. 6022-6033.
[3] W. Gao and S. Engell. Iterative Set-point Optimization of Batch Chromatography, Comp. Chem. Eng., 29 (2005), pp. 1401-1409.
[4] S. Skogestad. Plantwide Control: The Search for the Self-Optimizing Control Structure, J. of Process Cont., 10 (2000), pp. 487-507.
[5] G. Francois, B. Srinivasan and D. Bonvin. Use of Measurements for Enforcing the Necessary Conditions of Optimality in the Presence of Uncertainty and Constraints, J. of Process Cont., 15, n. 6 (2005), pp. 701-712.
[6] S. Boyd and L. Vandenberghe. Convex Optimization; Cambridge University Press, 2004.
[7] O. Rotava and A.C. Zanin. Multivariable control and real-time optimization – An industrial practical view. Hydrocarbon Processing 84 (2005), pp. 61-71.
[8] F.J. Serralunga, M.C. Mussati and P.A. Aguirre. Model Adaptation for Real-Time Optimization in Energy Systems. Ind. Eng. Chem. Res., 52 (2013), pp 16795-16810.
[9] M. Bazarra, H. Sherali, and C. Shetty, Nonlinear Programming: Theory and Algorithms. John Wiley and Sons, 2nd edition: New York, USA (1993).
[10] J.F. Forbes and T. Marlin, T. Design Cost: A systematic approach to technology selection for model-based real-time optimization systems. Comput. Chem. Eng., 20 (1996), pp. 717–734.
[11] G. Francois and D. Bonvin. Use of Convex Model Approximations for Real-Time Optimization via Modifier Adaptation, Ind. Eng, Chem. Res., 52, num. 33, (2013), pp. 11614-11625.
[12] G. Bunin, G. Francois and D. Bonvin. From Discrete Measurements to Bounded Gradient Estimates: A Look at Some Regularizing Structures, Ind. Eng. Chem. Res., 52, num. 35, (2013), pp. 12500-12513.
[13] G. Francois, B. Srinivasan and D. Bonvin. Comparison of Six Implicit Real-Time Optimization Schemes, Journal Européen des Systèmes Automatisés, 46, num. 2-3, (2012), pp. 291-305.
[14] B. Srinivasan, G. Francois and D. Bonvin. Comparison of Gradient Estimation Methods for Real-time Optimization, Computer Aided Chemical Engineering, 29 (2011), pp. 607-611.
[15] G.A. Korn and T. M. Korn. Mathematical Handbook for Scientists and Engineers. Dover Publications (2000).
[16] G. Bunin, G. Francois and D. Bonvin. Feasible-side global convergence in experimental optimization, SIAM Journal on Optimization, submitted, (2014).
[17] G. Bunin, G. Francois, and D. Bonvin. The SCFO Real-Time Optimization Solver: Users' Guide (version 0.9.4). EPFL. http://infoscience.epfl.ch/record/186672, downloaded August 2013.
[18] S. Diethelm, J. Van Herle, Z. Wuillemin, A. Nakajo, N. Autissier and M. Molinelli. Impact of materials and design on solid oxide fuel cell stack operation. J. Fuel Cell Sci. and Tech., 5, num. 3, (2008), pp 1-6.
[19] Z. Wuillemin. Experimental and modeling investigations on local performance and local degradation in SOFC (no. 4525). Ph.D. thesis, EPFL, (2009).
[20] A. Marchetti. Modifier-adaptation methodology for real-time optimization (No. 4449). PhD thesis, EPFL, (2009).
[21] G. Bunin, Z. Wuillemin, G. Francois, A. Nakajo, L. Tsikonis and D. Bonvin. Experimental Real-Time Optimization of a Solid Oxide Fuel Cell Stack via Constraint Adaptation, Energy, 39, num. 1, (2012), pp. 54-62.
[22] A. Nakajo, Z. Wuillemin, J. Van Herle, J. and D. Favrat. Simulation of thermal stresses in anode-supportedsolid oxide fuel cell stacks. Part I : probability of failure of the cells, J. of Power Sources, 2913, num 1, (2009), pp. 203-215.
[23] M. Diehl. Real-Time Optimization for Large Scale Nonlinear Processes . Ph.D. thesis, Ruprecht-Karls- Universität Heidelberg, (2001).
[24] I. Argatov, P. Rautakorpi and R. Silvennoinen (2009). Estimation of the mechanical energy output of the kite wind generator. Renewable Energy, 34, num 6, (2009), pp 1525 - 1532.
[25] J. Breukels. An Engineering Methodology for Kite Design. Ph.D. thesis, Delft Universtiy of Technology, (2010).
[26] S. Costello, G. Francois and D. Bonvin, Real-Time Optimization for Kites. In Proceedings of the 5th IFAC International Workshop on Periodic Control Systems (2013), pp. 64-69.
[27] H. Maurer and C. Bükens. Sensitivity analysis and real-time control of parametric optimal control problems using boundary value methods. Online Optimization of Large Scale Systems, (2001), pp.17–56